

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/697,720  
Source: IFWO  
Date Processed by STIC: 05/11/2006

# ***ENTERED***



IFWO

## RAW SEQUENCE LISTING

DATE: 05/11/2006

PATENT APPLICATION: US/10/697,720A

TIME: 12:14:24

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\05112006\J697720A.raw

```

3 <110> APPLICANT: Cohen, Stanley N.
4     Li, Limin
6 <120> TITLE OF INVENTION: MAMMALIAN TUMOR SUSCEPTIBILITY GENES AND THEIR USES
8 <130> FILE REFERENCE: 70017.27USC2
10 <140> CURRENT APPLICATION NUMBER: US 10/697,720A
11 <141> CURRENT FILING DATE: 2003-10-29
13 <150> PRIOR APPLICATION NUMBER: US 09/804,690
14 <151> PRIOR FILING DATE: 2001-03-12
16 <150> PRIOR APPLICATION NUMBER: US 09/146,187
17 <151> PRIOR FILING DATE: 1998-09-01
19 <150> PRIOR APPLICATION NUMBER: US 08/977,818
20 <151> PRIOR FILING DATE: 1997-11-25
22 <150> PRIOR APPLICATION NUMBER: US 08/670,274
23 <151> PRIOR FILING DATE: 1996-06-13
25 <150> PRIOR APPLICATION NUMBER: US 08/585,758
26 <151> PRIOR FILING DATE: 1996-01-16
28 <150> PRIOR APPLICATION NUMBER: US 60/006,856
29 <151> PRIOR FILING DATE: 1995-11-16
31 <160> NUMBER OF SEQ ID NOS: 20
33 <170> SOFTWARE: PatentIn version 3.3
35 <210> SEQ ID NO: 1
36 <211> LENGTH: 1448
37 <212> TYPE: DNA
38 <213> ORGANISM: Artificial Sequence
40 <220> FEATURE:
41 <223> OTHER INFORMATION: TSG101 nucleotide
44 <220> FEATURE:
45 <221> NAME/KEY: CDS
46 <222> LOCATION: (61)..(1203)
48 <400> SEQUENCE: 1
49 cccctctgcc tgtggggacg gaggagcgcg ccatggctgt ccgagagtca gctgaagaag      60
51 atg atg tcc aag tac aaa tat aga gat cta acc gtc cgt caa act gtc      108
52 Met Met Ser Lys Tyr Lys Tyr Arg Asp Leu Thr Val Arg Gln Thr Val
53 1           5           10           15
55 aat gtc atc gct atg tac aaa gat ctc aaa cct gta ttg gat tca tat      156
56 Asn Val Ile Ala Met Tyr Lys Asp Leu Lys Pro Val Leu Asp Ser Tyr
57           20           25           30
59 gtt ttt aat gat ggc agt tcc agg gag ctg gtg aac ctc act ggt aca      204
60 Val Phe Asn Asp Gly Ser Ser Arg Glu Leu Val Asn Leu Thr Gly Thr
61           35           40           45
63 atc cca gtg cgt tat cga ggt aat ata tat aat att cca ata tgc ctg      252
64 Ile Pro Val Arg Tyr Arg Gly Asn Ile Tyr Asn Ile Pro Ile Cys Leu
65     50           55           60

```

## RAW SEQUENCE LISTING

DATE: 05/11/2006

PATENT APPLICATION: US/10/697,720A

TIME: 12:14:24

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\05112006\J697720A.raw

67	tgg	ctg	ctg	gac	aca	tac	cca	tat	aac	ccc	cct	atc	tgt	ttt	gtt	aag	300
68	Trp	Leu	Leu	Asp	Thr	Tyr	Pro	Tyr	Asn	Pro	Pro	Ile	Cys	Phe	Val	Lys	
69	65					70				75					80		
71	cct	act	agt	tca	atg	act	att	aaa	aca	gga	aag	cat	gtg	gat	gca	aat	348
72	Pro	Thr	Ser	Ser	Met	Thr	Ile	Lys	Thr	Gly	Lys	His	Val	Asp	Ala	Asn	
73					85					90					95		
75	ggg	aaa	atc	tac	cta	cct	tat	cta	cat	gac	tgg	aaa	cat	cca	cgg	tca	396
76	Gly	Lys	Ile	Tyr	Leu	Pro	Tyr	Leu	His	Asp	Trp	Lys	His	Pro	Arg	Ser	
77				100					105					110			
79	gag	ttg	ctg	gag	ctt	att	caa	atc	atg	att	gtg	ata	ttt	gga	gag	gag	444
80	Glu	Leu	Leu	Glu	Leu	Ile	Gln	Ile	Met	Ile	Val	Ile	Phe	Gly	Glu	Glu	
81			115					120					125				
83	cct	cca	gtg	ttc	tcc	cgg	cct	act	gtt	tct	gca	tcc	tac	cca	cca	tac	492
84	Pro	Pro	Val	Phe	Ser	Arg	Pro	Thr	Val	Ser	Ala	Ser	Tyr	Pro	Pro	Tyr	
85		130					135				140						
87	aca	gca	aca	ggg	cca	cca	aat	acc	tcc	tac	atg	cca	ggc	atg	cca	agt	540
88	Thr	Ala	Thr	Gly	Pro	Pro	Asn	Thr	Ser	Tyr	Met	Pro	Gly	Met	Pro	Ser	
89	145				150					155				160			
91	gga	atc	tct	gca	tat	cca	tct	gga	tac	cct	ccc	aac	ccc	agt	ggg	tat	588
92	Gly	Ile	Ser	Ala	Tyr	Pro	Ser	Gly	Tyr	Pro	Pro	Asn	Pro	Ser	Gly	Tyr	
93				165					170					175			
95	cct	ggc	tgt	cct	tac	cca	cct	gct	ggc	cca	tac	cct	gcc	aca	aca	agc	636
96	Pro	Gly	Cys	Pro	Tyr	Pro	Pro	Ala	Gly	Pro	Tyr	Pro	Ala	Thr	Thr	Ser	
97			180					185					190				
99	tca	cag	tac	cct	tcc	cag	cct	cct	gtg	acc	act	gtt	ggg	ccc	agc	aga	684
100	Ser	Gln	Tyr	Pro	Ser	Gln	Pro	Pro	Val	Thr	Thr	Val	Gly	Pro	Ser	Arg	
101		195					200					205					
103	gat	ggc	aca	atc	agt	gag	gac	act	atc	cgt	gca	tct	ctc	atc	tca	gca	732
104	Asp	Gly	Thr	Ile	Ser	Glu	Asp	Thr	Ile	Arg	Ala	Ser	Leu	Ile	Ser	Ala	
105		210				215					220						
107	gtc	agt	gac	aaa	ctg	aga	tgg	cgg	atg	aag	gag	gaa	atg	gat	ggg	gcc	780
108	Val	Ser	Asp	Lys	Leu	Arg	Trp	Arg	Met	Lys	Glu	Glu	Met	Asp	Gly	Ala	
109	225				230					235				240			
111	cag	gca	gag	ctt	aat	gcc	ttg	aaa	cga	aca	gag	gaa	gat	ctg	aaa	aaa	828
112	Gln	Ala	Glu	Leu	Asn	Ala	Leu	Lys	Arg	Thr	Glu	Glu	Asp	Leu	Lys	Lys	
113				245					250					255			
115	ggc	cac	cag	aaa	ctg	gaa	gag	atg	gtc	acc	cgc	tta	gat	caa	gaa	gta	876
116	Gly	His	Gln	Lys	Leu	Glu	Glu	Met	Val	Thr	Arg	Leu	Asp	Gln	Glu	Val	
117			260					265					270				
119	gct	gaa	gtt	gat	aaa	aac	ata	gaa	ctt	ttg	aaa	aag	aag	gat	gaa	gaa	924
120	Ala	Glu	Val	Asp	Lys	Asn	Ile	Glu	Leu	Leu	Lys	Lys	Lys	Asp	Glu	Glu	
121		275				280					285						
123	cta	agt	tct	gct	ctg	gag	aaa	atg	gaa	aat	caa	tct	gaa	aat	aat	gat	972
124	Leu	Ser	Ser	Ala	Leu	Glu	Lys	Met	Glu	Asn	Gln	Ser	Glu	Asn	Asn	Asp	
125		290				295					300						
127	att	gat	gaa	gtt	atc	att	ccc	aca	gcc	cca	ctg	tat	aaa	cag	att	cta	1020
128	Ile	Asp	Glu	Val	Ile	Ile	Pro	Thr	Ala	Pro	Leu	Tyr	Lys	Gln	Ile	Leu	
129	305				310					315				320			
131	aat	ctg	tat	gca	gag	gaa	aat	gct	att	gaa	gac	act	atc	ttt	tac	ctt	1068

## RAW SEQUENCE LISTING

DATE: 05/11/2006

PATENT APPLICATION: US/10/697,720A

TIME: 12:14:24

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\05112006\J697720A.raw

```

132 Asn Leu Tyr Ala Glu Glu Asn Ala Ile Glu Asp Thr Ile Phe Tyr Leu
133          325          330          335
135 gga gaa gct ttg cgg cgg gga gtc ata gac ctg gat gtg ttc ctg aaa      1116
136 Gly Glu Ala Leu Arg Arg Gly Val Ile Asp Leu Asp Val Phe Leu Lys
137          340          345          350
139 cac gtc cgc ctc ctg tcc cgt aaa cag ttc cag cta agg gca cta atg      1164
140 His Val Arg Leu Leu Ser Arg Lys Gln Phe Gln Leu Arg Ala Leu Met
141          355          360          365
143 caa aag gca agg aag act gcg ggc ctt agt gac ctc tac tgacatgtgc      1213
144 Gln Lys Ala Arg Lys Thr Ala Gly Leu Ser Asp Leu Tyr
145          370          375          380
147 tgtcagctgg agaccgacct ctccgtaaag cattcttttc ttcttctttt tctcatcagt      1273
149 agaaccacaca ataagttatt gcagtttatc attcaagtgt taaatatttt gaatcaataa      1333
151 tatattttct gtttcctttg ggtaaaaact ggctttttatt aatgcacttt ctaccctctg      1393
153 taagcgtctg tgctgtgctg ggactgactg ggctaaataa aatttggtgc ataaa      1448
156 <210> SEQ ID NO: 2
157 <211> LENGTH: 381
158 <212> TYPE: PRT
159 <213> ORGANISM: Artificial Sequence
161 <220> FEATURE:
162 <223> OTHER INFORMATION: TSG101 amino acid
164 <400> SEQUENCE: 2
166 Met Met Ser Lys Tyr Lys Tyr Arg Asp Leu Thr Val Arg Gln Thr Val
167 1          5          10          15
170 Asn Val Ile Ala Met Tyr Lys Asp Leu Lys Pro Val Leu Asp Ser Tyr
171          20          25          30
174 Val Phe Asn Asp Gly Ser Ser Arg Glu Leu Val Asn Leu Thr Gly Thr
175          35          40          45
178 Ile Pro Val Arg Tyr Arg Gly Asn Ile Tyr Asn Ile Pro Ile Cys Leu
179          50          55          60
182 Trp Leu Leu Asp Thr Tyr Pro Tyr Asn Pro Pro Ile Cys Phe Val Lys
183 65          70          75          80
186 Pro Thr Ser Ser Met Thr Ile Lys Thr Gly Lys His Val Asp Ala Asn
187          85          90          95
190 Gly Lys Ile Tyr Leu Pro Tyr Leu His Asp Trp Lys His Pro Arg Ser
191          100         105         110
194 Glu Leu Leu Glu Leu Ile Gln Ile Met Ile Val Ile Phe Gly Glu Glu
195          115         120         125
198 Pro Pro Val Phe Ser Arg Pro Thr Val Ser Ala Ser Tyr Pro Pro Tyr
199          130         135         140
202 Thr Ala Thr Gly Pro Pro Asn Thr Ser Tyr Met Pro Gly Met Pro Ser
203 145         150         155         160
206 Gly Ile Ser Ala Tyr Pro Ser Gly Tyr Pro Pro Asn Pro Ser Gly Tyr
207          165         170         175
210 Pro Gly Cys Pro Tyr Pro Pro Ala Gly Pro Tyr Pro Ala Thr Thr Ser
211          180         185         190
214 Ser Gln Tyr Pro Ser Gln Pro Pro Val Thr Thr Val Gly Pro Ser Arg
215          195         200         205
218 Asp Gly Thr Ile Ser Glu Asp Thr Ile Arg Ala Ser Leu Ile Ser Ala

```

## RAW SEQUENCE LISTING

DATE: 05/11/2006

PATENT APPLICATION: US/10/697,720A

TIME: 12:14:24

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\05112006\J697720A.raw

219	210	215	220				
222	Val Ser Asp Lys Leu Arg	Trp Arg Met Lys Glu	Glu Met Asp Gly Ala				
223	225	230	235	240			
226	Gln Ala Glu Leu Asn Ala	Leu Lys Arg Thr Glu	Glu Asp Leu Lys Lys				
227	245	250	255				
230	Gly His Gln Lys Leu Glu	Glu Met Val Thr Arg	Leu Asp Gln Glu Val				
231	260	265	270				
234	Ala Glu Val Asp Lys Asn	Ile Glu Leu Leu Lys	Lys Lys Asp Glu Glu				
235	275	280	285				
238	Leu Ser Ser Ala Leu Glu	Lys Met Glu Asn Gln	Ser Glu Asn Asn Asp				
239	290	295	300				
242	Ile Asp Glu Val Ile Ile	Pro Thr Ala Pro Leu	Tyr Lys Gln Ile Leu				
243	305	310	315	320			
246	Asn Leu Tyr Ala Glu Glu	Asn Ala Ile Glu Asp	Thr Ile Phe Tyr Leu				
247	325	330	335				
250	Gly Glu Ala Leu Arg Arg	Gly Val Ile Asp Leu	Asp Val Phe Leu Lys				
251	340	345	350				
254	His Val Arg Leu Leu Ser	Arg Lys Gln Phe Gln	Leu Arg Ala Leu Met				
255	355	360	365				
258	Gln Lys Ala Arg Lys Thr	Ala Gly Leu Ser Asp	Leu Tyr				
259	370	375	380				
262	<210> SEQ ID NO: 3						
263	<211> LENGTH: 1494						
264	<212> TYPE: DNA						
265	<213> ORGANISM: Artificial Sequence						
267	<220> FEATURE:						
268	<223> OTHER INFORMATION: Primer						
270	<400> SEQUENCE: 3						
271	gaagggtgtg	cgattgtgtg	ggacggtctg	gggcagccca	gcagcggctg	accctctgcc	60
273	tgcggggaag	ggagtcgcca	ggcggccgtc	atggcgggtg	cggagagcca	gctcaagaaa	120
275	atggtgtcca	agtacaaata	cagagacctt	actgtacgtg	aaactgtcaa	tggtattact	180
277	ctatacaaag	atctcaaacc	tgttttggat	tcatatgttt	ttaacgatgg	cagttccagg	240
279	gaactaatga	acctcactgg	aacaatccct	gtgccttata	gaggttaatac	atacaatatt	300
281	ccaatatgcc	tatggctact	ggacacatac	ccatataatc	cccctatctg	ttttgttaag	360
283	cctactagtt	caatgactat	taaaacagga	aagcatgttg	atgcaaattg	gaagatatat	420
285	cttccttata	tacatgaatg	gaaacaccca	cagtcagact	tggtggggct	tattcagggtc	480
287	atgattgtgg	tatttggaga	tgaacctcca	gtcttctctc	gtcctatttc	ggcctcctat	540
289	ccgccatacc	aggcaacggg	gccaccaa	acttcttaca	tgccaggcat	gccagggtga	600
291	atctctccat	acccatccgg	ataccctccc	aatcccagtg	gttaccagag	ctgtccttac	660
293	ccacctggtg	gtccatatcc	tgccacaaca	agttctcagt	acccttctca	gcctcctgtg	720
295	accactgttg	gtcccagtag	ggatggcaca	atcagcgagg	acaccatccg	agcctctctc	780
297	atctctgcgg	tcagtgacaa	actgagatgg	cggatgaagg	aggaaatgga	tcgtgccag	840
299	gcagagctca	atgccttgaa	acgaacagaa	gaagacctga	aaaagggtca	ccagaaactg	900
301	gaagagatgg	ttaccggttt	agatcaagaa	gtagccgagg	ttgataaaaa	catagaactt	960
303	ttgaaaaaga	aggatgaaga	actcagttct	gctctggaaa	aaatggaaaa	tcagtctgaa	1020
305	aacaatgata	tcgatgaagt	tatcattccc	acagctccct	tatacaaa	gatcctgaat	1080
307	ctgtatgcag	aagaaaacgc	tattgaagac	actatctttt	acttgggaga	agccttgaga	1140
309	aggggcgtga	tagacctgga	tgtcttctct	aagcatgtac	gtcttctgtc	ccgtaaacag	1200
311	ttccagctga	gggcactaat	gcaaaaagca	agaaagactg	ccggtctcag	tgacctctac	1260

## RAW SEQUENCE LISTING

DATE: 05/11/2006

PATENT APPLICATION: US/10/697,720A

TIME: 12:14:24

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\05112006\J697720A.raw

```

313 tgactttctct gataccagct ggaggttgag ctcttcttaa agtattcttc tcttcctttt 1320
315 atcagtaggt gccagaata agttattgca gtttatcatt caagtgtaaa atattttgaa 1380
317 tcaataatat attttctgtt ttcttttggg aaagactggc ttttattaat gcactttcta 1440
319 tcctctgtaa actttttgtg ctgaatgttg ggactgctaa ataaaatttg tttt 1494
322 <210> SEQ ID NO: 4
323 <211> LENGTH: 390
324 <212> TYPE: PRT
325 <213> ORGANISM: Artificial Sequence
327 <220> FEATURE:
328 <223> OTHER INFORMATION: Primer
330 <400> SEQUENCE: 4
332 Met Ala Val Ser Glu Ser Gln Leu Lys Lys Met Val Ser Lys Tyr Lys
333 1 5 10 15
336 Tyr Arg Asp Leu Thr Val Arg Glu Thr Val Asn Val Ile Thr Leu Tyr
337 20 25 30
340 Lys Asp Leu Lys Pro Val Leu Asp Ser Tyr Val Phe Asn Asp Gly Ser
341 35 40 45
344 Ser Arg Glu Leu Met Asn Leu Thr Gly Thr Ile Pro Val Pro Tyr Arg
345 50 55 60
348 Gly Asn Thr Tyr Asn Ile Pro Ile Cys Leu Trp Leu Leu Asp Thr Tyr
349 65 70 75 80
352 Pro Tyr Asn Pro Pro Ile Cys Phe Val Lys Pro Thr Ser Ser Met Thr
353 85 90 95
356 Ile Lys Thr Gly Lys His Val Asp Ala Asn Gly Lys Ile Tyr Leu Pro
357 100 105 110
360 Tyr Leu His Glu Trp Lys His Pro Gln Ser Asp Leu Leu Gly Leu Ile
361 115 120 125
364 Gln Val Met Ile Val Val Phe Gly Asp Glu Pro Pro Val Phe Ser Arg
365 130 135 140
368 Pro Ile Ser Ala Ser Tyr Pro Pro Tyr Gln Ala Thr Gly Pro Pro Asn
369 145 150 155 160
372 Thr Ser Tyr Met Pro Gly Met Pro Gly Gly Ile Ser Pro Tyr Pro Ser
373 165 170 175
376 Gly Tyr Pro Pro Asn Pro Ser Gly Tyr Pro Gly Cys Pro Tyr Pro Pro
377 180 185 190
380 Gly Gly Pro Tyr Pro Ala Thr Thr Ser Ser Gln Tyr Pro Ser Gln Pro
381 195 200 205
384 Pro Val Thr Thr Val Gly Pro Ser Arg Asp Gly Thr Ile Ser Glu Asp
385 210 215 220
388 Thr Ile Arg Ala Ser Leu Ile Ser Ala Val Ser Asp Lys Leu Arg Trp
389 225 230 235 240
392 Arg Met Lys Glu Glu Met Asp Arg Ala Gln Ala Glu Leu Asn Ala Leu
393 245 250 255
396 Lys Arg Thr Glu Glu Asp Leu Lys Lys Gly His Gln Lys Leu Glu Glu
397 260 265 270
400 Met Val Thr Arg Leu Asp Gln Glu Val Ala Glu Val Asp Lys Asn Ile
401 275 280 285
404 Glu Leu Leu Lys Lys Lys Asp Glu Glu Leu Ser Ser Ala Leu Glu Lys
405 290 295 300

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/697,720A

DATE: 05/11/2006  
TIME: 12:14:25

Input Set : A:\Sequence Listing.txt  
Output Set: N:\CRF4\05112006\J697720A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:20; Xaa Pos. 2

**VERIFICATION SUMMARY**

DATE: 05/11/2006

PATENT APPLICATION: US/10/697,720A

TIME: 12:14:25

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\05112006\J697720A.raw

L:630 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20 after pos.:0